ISSN 2347-1921

ON PARAMETRIC ESTIMATION

Luela Prifti
Department of Mathematical Engineering, Polytechnic University of Tirana
Sheshi "Nene Tereza", nr. 1, Tirana, Albania.
luela_p@yahoo.com
Shpetim Shehu
Department of Mathematical Engineering, Polytechnic University of Tirana
Sheshi "Nene Tereza", nr. 1, Tirana, Albania.
shpetimi49@yahoo.com

ABSTRACT

In this paper we are focused on parametric modeling. To estimate the distribution of the population under study we use the point estimate. For the unknown parameter *a* of the uniform (continue) distribution U(0,*a*) we present a better estimator than the sample mean. It will provide a basis for developing efficient estimators when faced with similar problems.

Keywords

parametric estimator; uniform distribution.

Academic Discipline And Sub-Disciplines

Statistics;

SUBJECT CLASSIFICATION

Statistical subject classification

TYPE (METHOD/APPROACH)

Theorical

Council for Innovative Research

Peer Review Research Publishing System

Journal: JOURNAL OF ADVANCES IN MATHEMATICS

Vol .10, No.7

www.cirjam.com , editorjam@gmail.com

ISSN 2347-1921

INTRODUCTION

In statistical modeling to estimate the distribution of the population under study there are two approaches: parametric and nonparametric. For the parametric estimation we first need to choose a family of models within which the true distribution of data is supposed to lie and then to choose the best estimator for the parameters of the distribution. The statisticians present the parameter estimation process from different viewpoints focusing on the theory aspects [1] or in the applications [2]. In the first part of this paper we present the main ideas in estimation theory from a frequentist viewpoint focused on point estimator [3],[4] and in the second one for the uniform distribution U(0,*a*) an estimator more efficient than the sample mean is presented. It will provide a basis for developing efficient estimators when faced with similar problems.

1. BASIS OF PARAMETRIC ESTIMATION

Let $x_1, x_2, \ldots, x_n$ be independent realizations from a population distribution. To estimate this distribution we need to choose a model based on the data under study (using probability graph, or box plot). Generally for large *n* we can use the central limit theorem for the distribution of the sample mean.

Theorem

Let $X_1, X_2, \ldots, X_n$ be a sequence of independent and identically distributed random variables with mean μ and finite positive variance σ^2 then the sample mean of the above variables sequence $\overline{X}$ defined as:

$$\overline{X} = \frac{X_1 + X_2 + \ldots + X_n}{n}$$

is approximately distributed as normal distribution with parameter μ and $\frac{\sigma^2}{n}$.

Let *X* be a random variable with probability density function belonging to a known family of probability distributions with density functions

$$F = \{(f(x, \theta) : \theta \in \Theta)\}$$

and $X_1, X_2, \ldots, X_n$ be a sample for this random variable.

To estimate the true value of parameter θ denoted by θ_0 we use the estimator $\hat{\theta}_0$ which should be close to the parameter value θ_0. For this the estimator $\hat{\theta}_0$ should fulfill these conditions:

a) The estimator $\hat{\theta}_0$ should be with bias zero. It means that for it the value on average is the true parameter value.

$$E(\hat{\theta}_0) = \theta_0$$

b) The variation of the estimator around the true parameter value should be small.

$$Var(\hat{\theta}_0) = E(\hat{\theta}_0 - \theta_0)^2$$

The standard deviation of the $\hat{\theta}_0$ obtained from the data measures how precise the estimator is: the smaller the standard deviation of the $\hat{\theta}_0$ the greater the precision. Calculating a confidence interval for θ_0 we make it more explicit.

There are many methods for estimates unknown parameters from data. The estimator will involve intervals and probabilities.

We will consider the maximum likelihood estimator (MLE) as a point estimator, which in most of the cases is easy to compute and in the simplest cases match with our intuition.

For the sample $x_1, x_2, \ldots, x_n$ of the random variable X having the probability density function $f(x, \theta_0)$, the likelihood function is:

$$L(\theta) = \prod_{i=1}^{n} f(x_i, \theta)$$

The maximum likelihood estimator $\hat{\theta}_0$ of θ_0 is defined as the value of θ that maxims the likelihood function. Since the logarithm function is monotonic, the log-likelihood takes its maximum at the same point as the likelihood function, so we can obtain the maximum likelihood estimator by differentiating the log-likelihood and equating to zero.

2. ESTIMATION OF THE PARAMETER FOR THE UNIFORM (CONTINUE) DISTRIBUTION

Let X be a random variable from the population with uniform distribution U(0,*a*) the density is $\frac{1}{a}$ on $[0, a]$ and 0 otherwise.

Let $x_1, x_2, \ldots x_n$ be a sample for this random variable drawn from the above population. We want to estimate the parameter *a*. The most natural estimator is the sample mean. We define it as $\bar{X}$.

So, $\bar{X} = \frac{a}{2}$.

Based on $\bar{X}$ the first unbiased estimator for *a* is:

$\hat{\theta}_1 = 2\overline{X}$ with variance $V(\hat{\theta}_1) = \frac{a^2}{3n}$.

Using the MLE for the above X and $x_1, x_2, \ldots x_n$ the likelihood function is :

$$f(x_1, x_2, \ldots x_n / a) = (\frac{1}{a})^n$$

This is maximazied by making *a* as small as possible. The interval $[0, a]$ must include all the data. Thus the MLE for *a* [5] is:

$$\hat{\theta}_2 = max\{x_1, x_2, \ldots x_n\}.$$

The second estimator obtained from MLE is asymptotically unbiased and has asymptotically minimal variance:

$$E(\hat{\theta}_2) = \frac{na}{n+1}$$

$$Var(\hat{\theta}_2) = \frac{na^2}{(n+1)^2(n+2)}$$

We present the third estimator for the parameter *a*, as an unbiased estimator combining the second one.

$$\hat{\theta}_3 = \frac{n+1}{n}\hat{\theta}_2 ,$$

$$E(\hat{\theta}_3) = a .$$

The variance for $\hat{\theta}_3$ is:

$$Var(\hat{\theta}_3) = Var(\frac{(n+1)\hat{\theta}_2}{n})$$

$$Var(\hat{\theta}_3) = \frac{a^2}{n(n+2)}$$

$$Var(\hat{\theta}_3) \langle Var(\hat{\theta}_1) .$$

So the, $\hat{\theta}_3$ is more efficient than $\hat{\theta}_1$. As result $\hat{\theta}_3$ is a better estimator than $\hat{\theta}_1$ for the parameter *a.*

CONCLUSION

The estimation of the parameters of a statistical model is a very important process in statistics. In this paper we are focused on point estimator. For the unknown parameter *a* of the uniform distribution U (0,*a*) we present a better estimator than the sample mean. It will provide a basis for developing efficient estimators when faced with similar problems.

REFERENCES

[1] Casella, G and Berger, R.L. (1990): Statistical Inference, Pacific Grove: Wadsworth & Brooks/Cole.

[2] Hahn,G.J. and Meeker, W.O.(1991): Statistical intervals- A guide for practitioners, New York :Wiley.

[3] Lehmann, E.L.(1993). Theory of point estimation, New York: Wiley.

[4] Mood, A.M., Graybill, F.A.and Boes, D.C. (1974). Introduction to the Theory of Statistics, New York: McGraw-Hill.

[5] Prifti, L., Shehu, Sh. (2012) Modelimi Statistikor i vlerave ekstreme. UNIEL,Scientific Buletin 2012 /1.16-29.